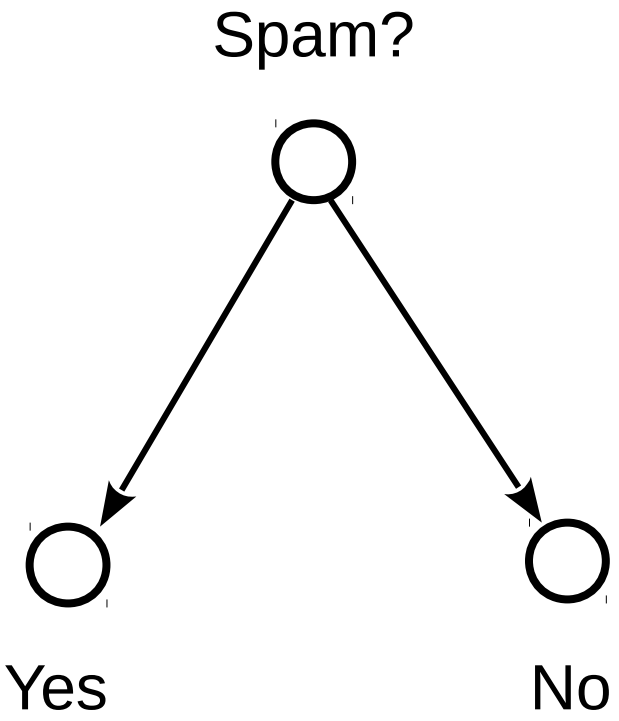
Commodity Machine Learning

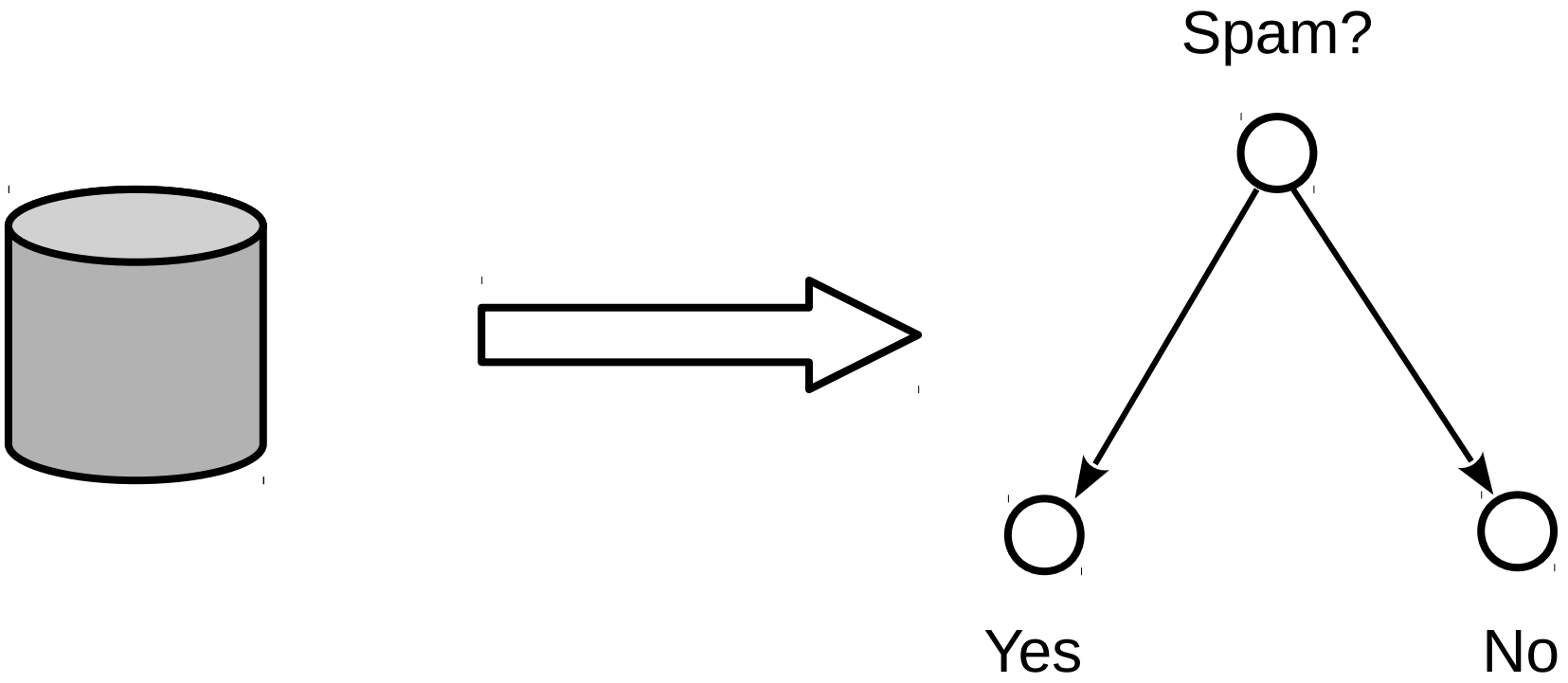
Past, present and future

Andreas Mueller

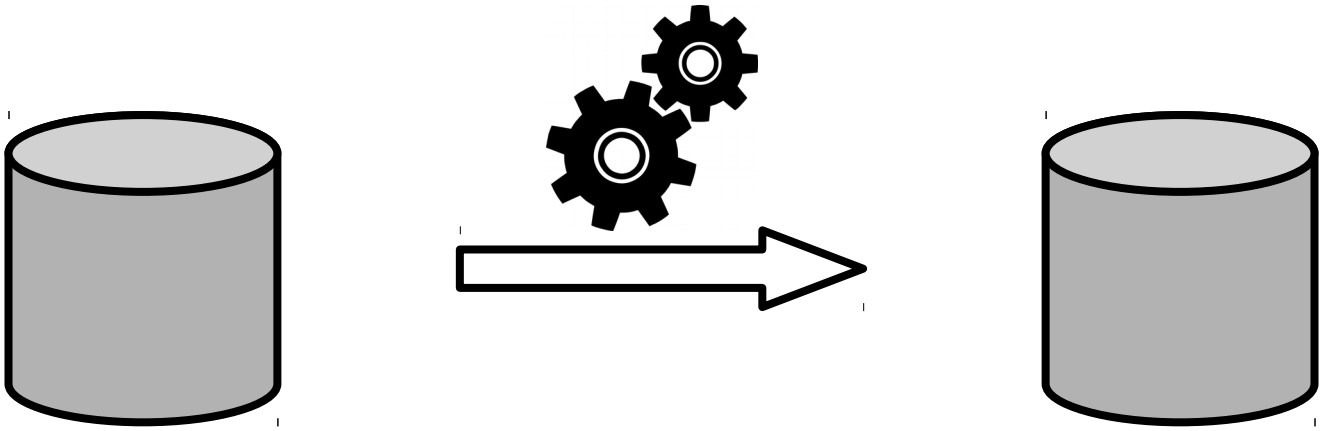
What is machine learning?

Automatic Decision Making

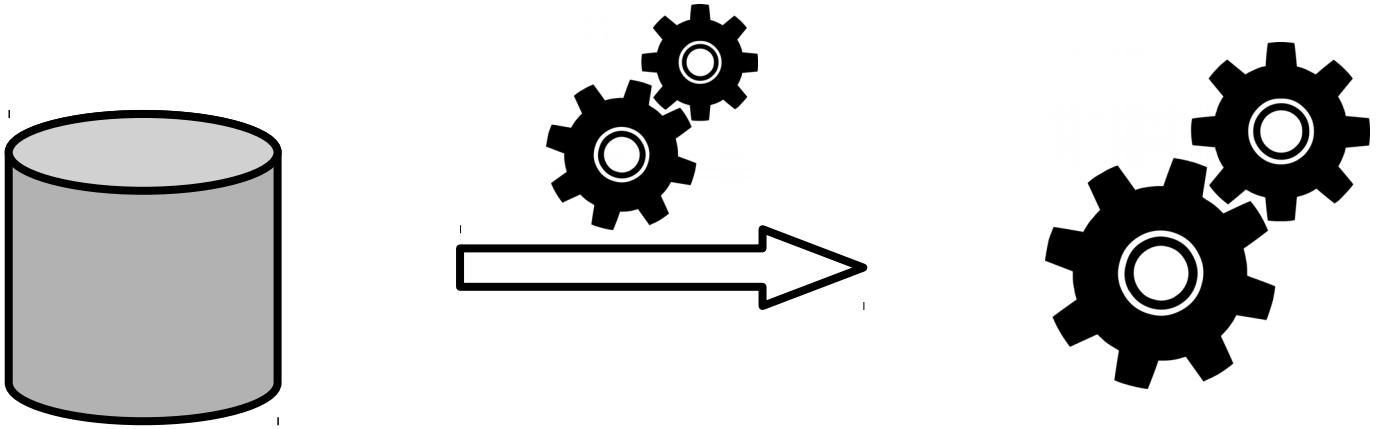



Spam?
Yes
No

Programming



Machine Learning



Machine learning is EVERYWHERE

Search for people, places and things

Andreas | Home

Andreas Mueller
Edit Profile

- **News Feed**
- Messages
- Events 1

GROUPS

- Amazon Machine L... 5
- Surprise 6
- Tobis Geschenk 10
- VERSCHENK'S - ... 20+
- Hochzeitsgeschenk 6
- Mein nächster großer ...
- Berlin Startup Jobs 20+
- Create Group

APPS

- Games
- Glassdoor
- Photos
- Pedagogy
- Music
- Cities I've Visited
- On This Day
- Games Feed 20+

FRIENDS

- Amazon.com 20+
- Rudolf-Koch-Schule 5
- Uni Bonn 20+
- University of Bonn 20+
- Close Friends 20+
- Berlin, Germany A... 20+

INTERESTS

- Pages and Public Fig...

PAGES

Update Status | Add Photos/Video

What's on your mind?

[redacted]
21 mins ·

Eine aktuelle Broodwar Partie: Jaedong vs. Stork
Ab 12:21.30 http://www.twitch.tv/taketv/b/536234505

Like · Comment · Share

Write a comment...

[redacted]
12 hrs · Twitter ·

Evening at the #biergarten 27° in the evening, #Imisssummerinthecontinent http://t.co/r0HeaXplaX

Instagram
[redacted] on Instagram
INSTAGRAM

Like · Comment · Share · [redacted]

[redacted]kes this.

Write a comment...

CrossFit Icke Outdoo... tomorrow

SPONSORED See All

Online Essen bestellen
lieferheld.de

Neu mit Lieferheld: PLZ eingeben, Lieferdienst finden und genießen!

Globaler Chauffeurservice
blacklane.com

Fahren Sie eine Klasse besser zu günstigen Preisen – Blacklane ist weltweit verfügbar!

28,711 people like this

Das Samsung Galaxy S5
o2online.de

Jetzt inklusive Daten-Roaming und Allnet-Flat. Gleich online sichern.

Nikon S9700 mit 30x ZOOM
nikon.de

Immer näher dran mit Superzoom, Wi-Fi & GPS. Perfekte Urlaubsfotos sofort teilen.

4 Bands.Open Air.For Free

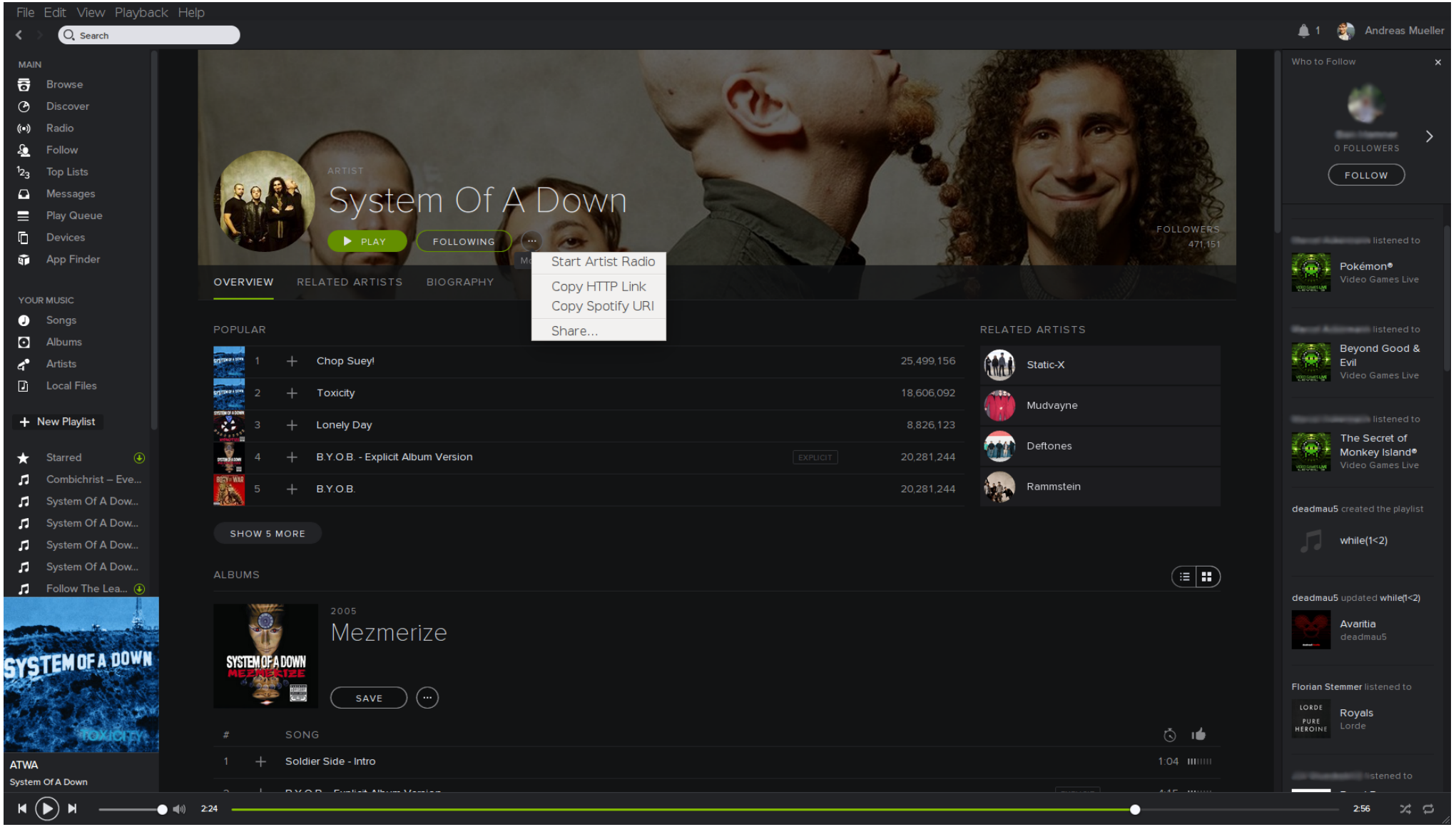
Mit Johannes Oerding u.v.m. EINTRITT FREI. 14.6. 18-22h direkt vor der O2 World Berlin.

Saturday, June 14 at 6:00pm

Join · 51 people are going

Singles auf Facebook

Schau dir Dating-Profile von Singles in deiner Nähe an.


File Edit View Playback Help
Search
1
Andreas Mueller
MAIN
Browse
Discover
Radio
Follow
Top Lists
Messages
Play Queue
Devices
App Finder
YOUR MUSIC
Songs
Albums
Artists
Local Files
New Playlist
Starred
Combichrist – Eve...
System Of A Dow...
System Of A Dow...
System Of A Dow...
System Of A Dow...
Follow The Lea...
ATWA
System Of A Down
ARTIST
System Of A Down
PLAY
FOLLOWING
FOLLOWERS
471,151
Start Artist Radio
Copy HTTP Link
Copy Spotify URI
Share...
OVERVIEW
RELATED ARTISTS
BIOGRAPHY
POPULAR
1 Chop Suey! 25,499,156
2 Toxicity 18,606,092
3 Lonely Day 8,826,123
4 B.Y.O.B. - Explicit Album Version EXPLICIT 20,281,244
5 B.Y.O.B. 20,281,244
SHOW 5 MORE
RELATED ARTISTS
Static-X
Mudvayne
Deftones
Rammstein
ALBUMS
2005
Mezmerize
SAVE
SONG
1 Soldier Side - Intro 1:04
Who to Follow
0 FOLLOWERS
FOLLOW
listened to
Pokémon®
Video Games Live
listened to
Beyond Good & Evil
Video Games Live
listened to
The Secret of Monkey Island®
Video Games Live
deadmau5 created the playlist
while(1<2)
deadmau5 updated while(1<2)
Avaritia
deadmau5
Florian Stemmer listened to
Royals
Lorde
2:24
2:56

www.amazon.com/s/ref=nb_sb_noss_1?url=search-alias%3Daps&field-keywords=machine learning&sprefix=machine+%2Caps&rh=

DuckDuckGo

amazon Try Prime | Andreas's Amazon.com | Today's Deals | Gift Cards | Sell | Help

Father's Day Is June 15 — Sponsored by DeWalt — Shop now

Shop by Department | Search All | machine learning | Go | Hello, Andreas Your Account | Try Prime | Cart | Wish List

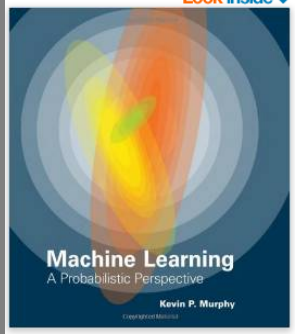
1-16 of 21,549 results for **"machine learning"** | Choose a Department to sort

Show results for

Books
- Machine Learning
- Computers & Technology

LOOK INSIDE! **Machine Learning Resources**
Find tips and tricks with these featured titles on machine learning, algorithms, sensors, and more. Learn more

Look inside



See all 1 image(s)

Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series) Hardcover

by Kevin P. Murphy (Author)

★★★★☆ 31 customer reviews

See all 2 formats and editions

Kindle	Hardcover
$64.06	$77.41
Read with our free app	20 Used from $67.99 42 New from $69.66

Today's Web-enabled deluge of electronic data calls for automated methods of data analysis. Machine learning provides these, developing methods that can automatically detect patterns in data and then use the uncovered patterns to predict future data. This textbook offers a comprehensive and self-contained introduction to the field of machine learning, based on a unified, probabilistic approach. The coverage combines breadth and depth, offering necessary background material on such topics as probability, optimization, and linear algebra as well as discussion of recent developments in the field, including conditional random fields, L1 regularization, and deep learning. The book is written in an informal, accessible style, complete with pseudo-code for the most important algorithms. All topics are copiously illustrated with color images and worked examples drawn from such application domains as

Read more

Share

Buy New $77.41

Qty: 1 | List Price: $90.00 | Save: $12.59 (14%)

FREE Shipping

In Stock.

Ships from and sold by Amazon.com. Gift-wrap available.

Add to Cart

Sign in to turn on 1-click ordering

Want it Wednesday, June 11? Order within **11 hrs 27 mins** and choose **AmazonGlobal Priority Shipping** at checkout. Details

Buy Used $71.98

Add to Wish List

Trade in your item Get a **$35.10** Gift Card. Trade in | Learn More

Have one to sell? Sell on Amazon

Frequently Bought Together

Price for all three: **$249.90**

Add all three to Cart | Add all three to Wish List

Show availability and shipping details

- **This item:** Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series) by Kevin P. Murphy Hardcover $77.41
- Pattern Recognition and Machine Learning (Information Science and Statistics) by Christopher M. Bishop Hardcover $85.46
- Probabilistic Graphical Models: Principles and Techniques (Adaptive Computation and Machine Learning ... by Daphne Koller Hardcover $87.03

Customers Who Bought This Item Also Bought

Page 1 of 12

Learning From Data
› Yaser S. Abu-Mostafa

Pattern Recognition and Machine Learning ...

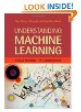
Probabilistic Graphical Models: Principles ...

The Elements of Statistical Learning: ...

Foundations of Machine Learning (Adaptive ...

Machine Learning: The Art and Science of ...

Building Machine Learning Systems with ...


Understanding Machine Learning: From Theory to

An Introduction to Statistical Learning: ...

Science
Engineering
Medicine
...

Commodity machine learning

past

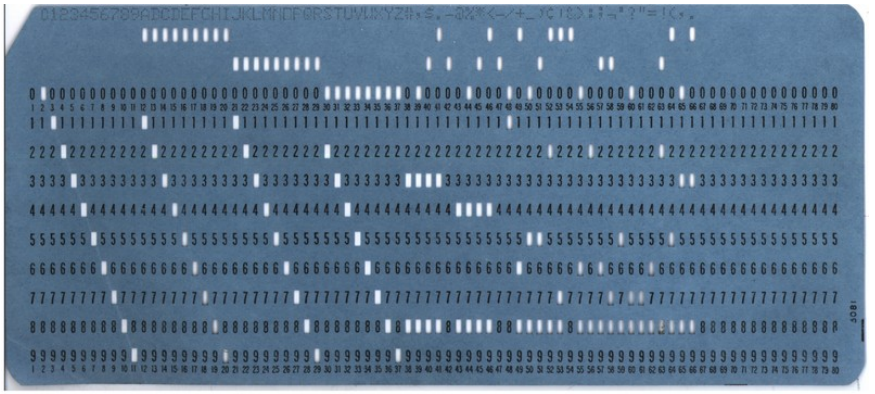
PHILOSOPHIÆ
NATURALIS
PRINCIPIA
MATHEMATICA.

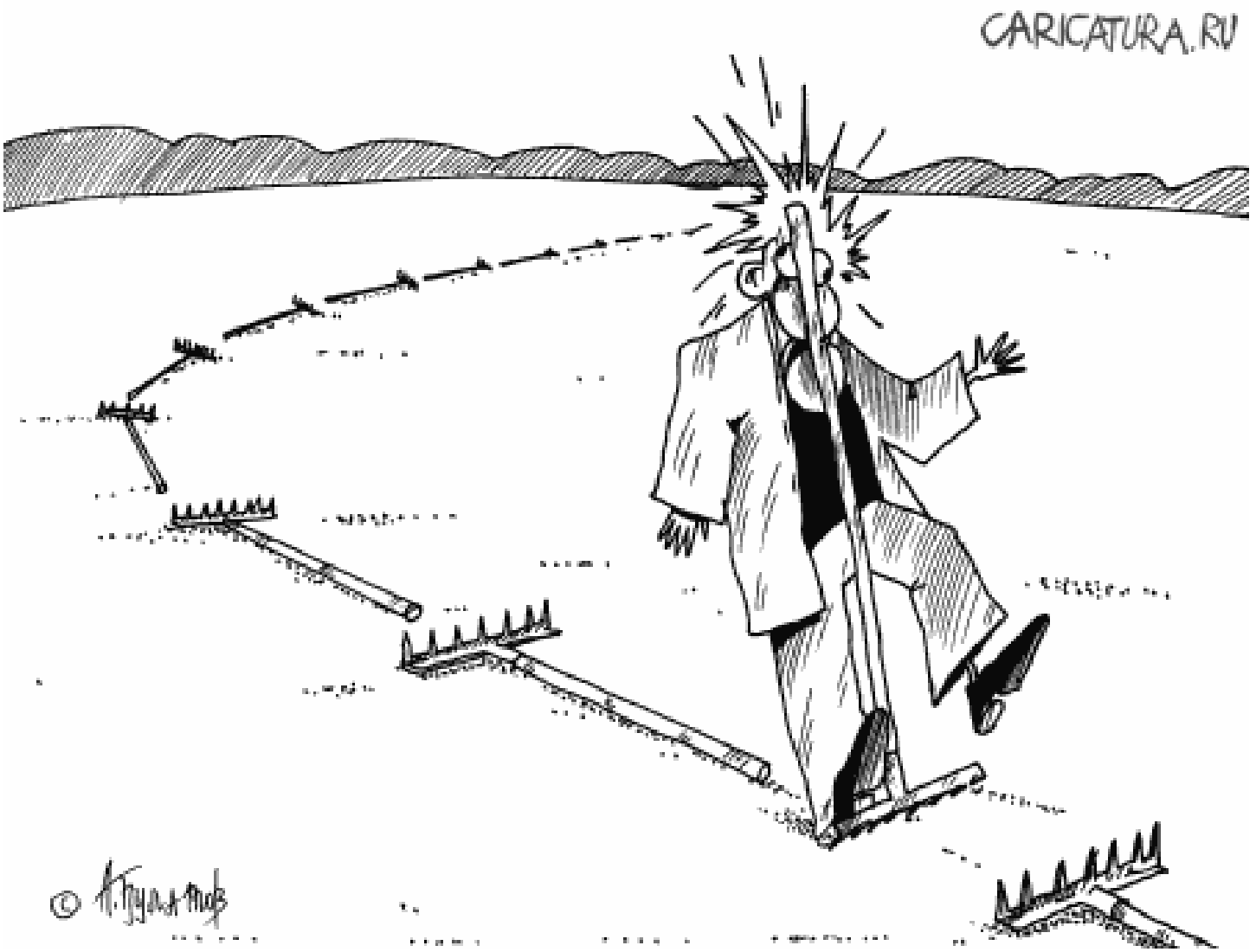
Autore *J S. NEWTON, Trin. Coll. Cantab. Soc.* Matheseos Professore *Lucasiano*, & Societatis Regalis Sodali.

IMPRIMATUR.
S. PEPYS, *Reg. Soc.* PRÆSES.
Julii 5. 1686.

LONDINI,
Jussu *Societatis Regiæ* ac Typis *Josephi Streater.* Prostat apud plures Bibliopolas. *Anno* MDCLXXXVII.

+




CARICATURA.RU

dawn of open source tools...

The age of shell

Documentation? Testing?

Scikit-learn: User centric machine learning

```
.fit(X, y)
.predict(X)
.transform(X)
```

present

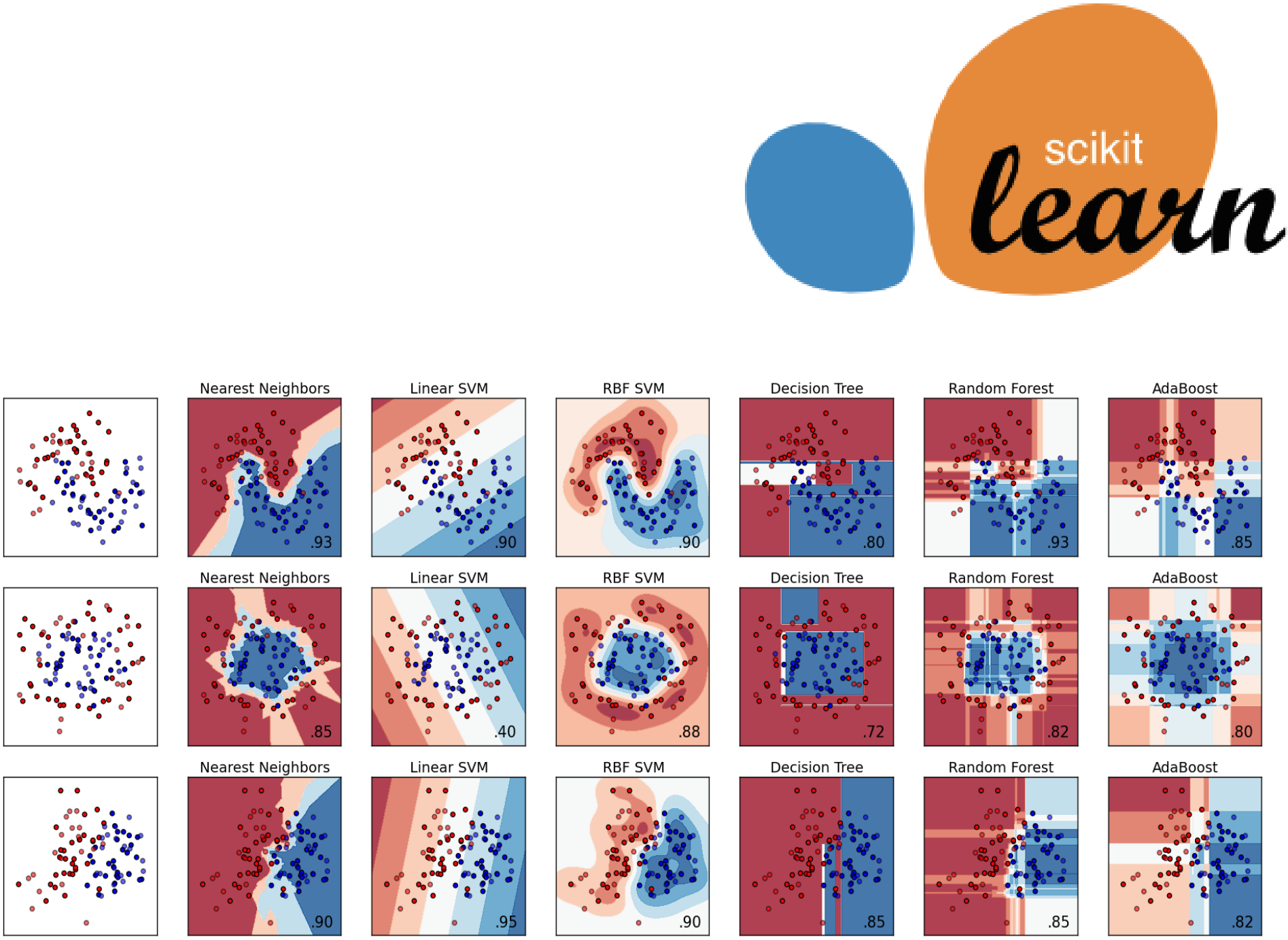
Choose your ecosystem.

Open! Documented! Tested!

Usability is key!

ML Frameworks

PyMC, Edward, Stan
theano, tensorflow, keras


Nearest Neighbors
Linear SVM
RBF SVM
Decision Tree
Random Forest
AdaBoost
.93
.90
.90
.80
.93
.85
Nearest Neighbors
Linear SVM
RBF SVM
Decision Tree
Random Forest
AdaBoost
.85
.40
.88
.72
.82
.80
Nearest Neighbors
Linear SVM
RBF SVM
Decision Tree
Random Forest
AdaBoost
.90
.95
.90
.85
.85
.82

```
from sklearn.model_selection import GridSearchCV
from sklearn.pipeline import Pipeline
```

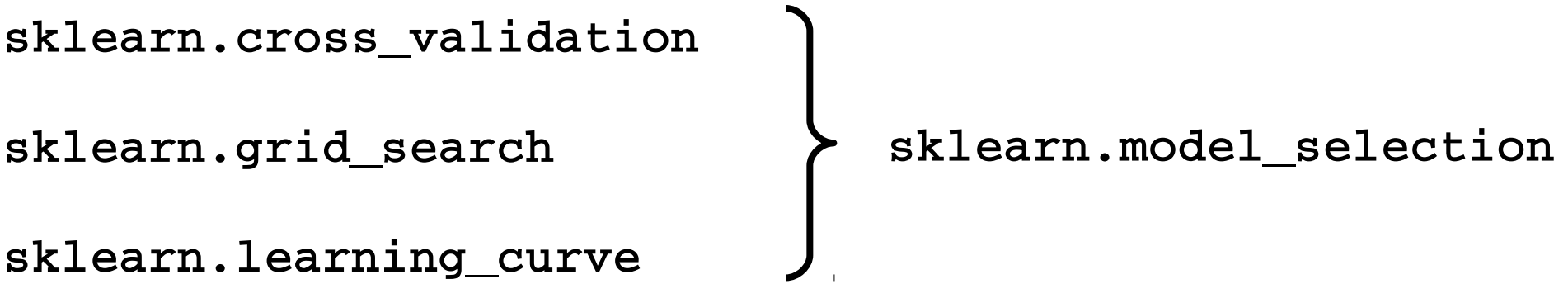
github.com/scikit-learn-contrib/scikit-learn-contrib

(near) Future

0.18

for the release candidate:

```
pip install scikit-learn==0.18rc2
```


sklearn.cross_validation
sklearn.grid_search
sklearn.learning_curve
sklearn.model_selection

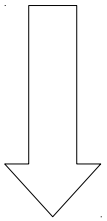
```
results = pd.DataFrame(grid_search.results_)
```

	mean_test_score	param_C	param_gamma	params	...	split2_test_score	split3_test_score	split4_test_score	std_test_score
0	0.37	0.001	0.001	{'C': 0.001, 'gamma': 0.001}	...	0.36	0.36	0.38	0.01
1	0.37	0.001	0.01	{'C': 0.001, 'gamma': 0.01}	...	0.36	0.36	0.38	0.01
2	0.37	0.001	0.1	{'C': 0.001, 'gamma': 0.1}	...	0.36	0.36	0.38	0.01
3	0.37	0.001	1	{'C': 0.001, 'gamma': 1}	...	0.36	0.36	0.38	0.01
4	0.37	0.001	10	{'C': 0.001, 'gamma': 10}	...	0.36	0.36	0.38	0.01

labels → groups

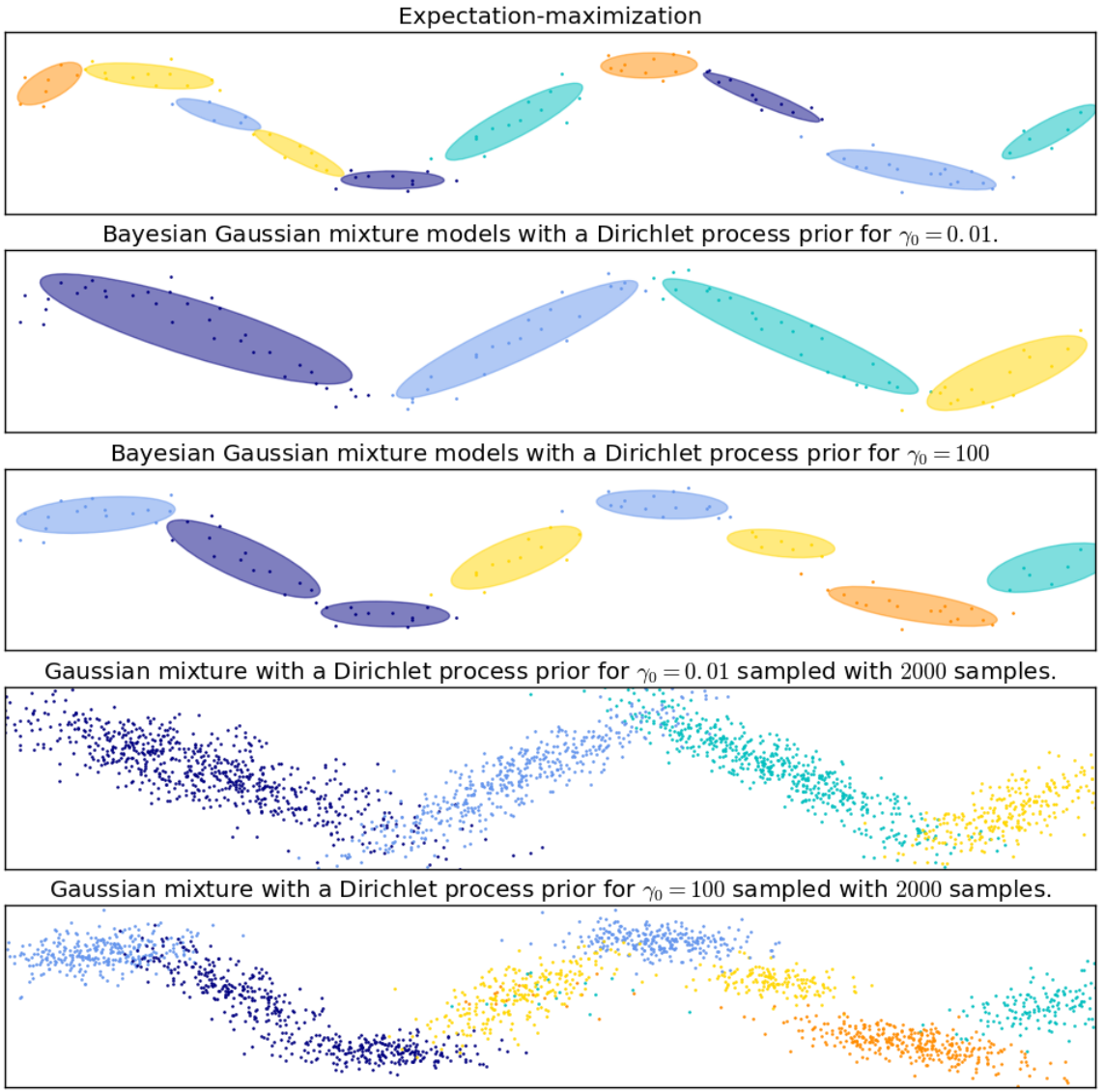
n_folds → n_splits

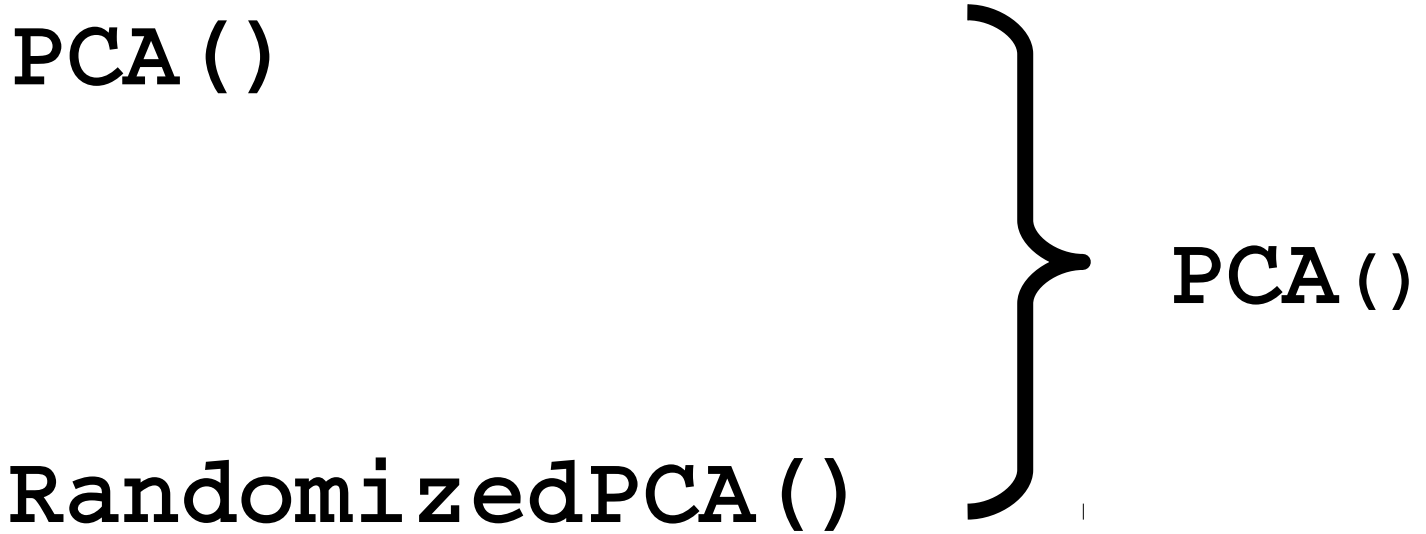
```
from sklearn.cross_validation import KFold
cv = KFold(n_samples, n_folds)
for train, test in cv:
    ...
```



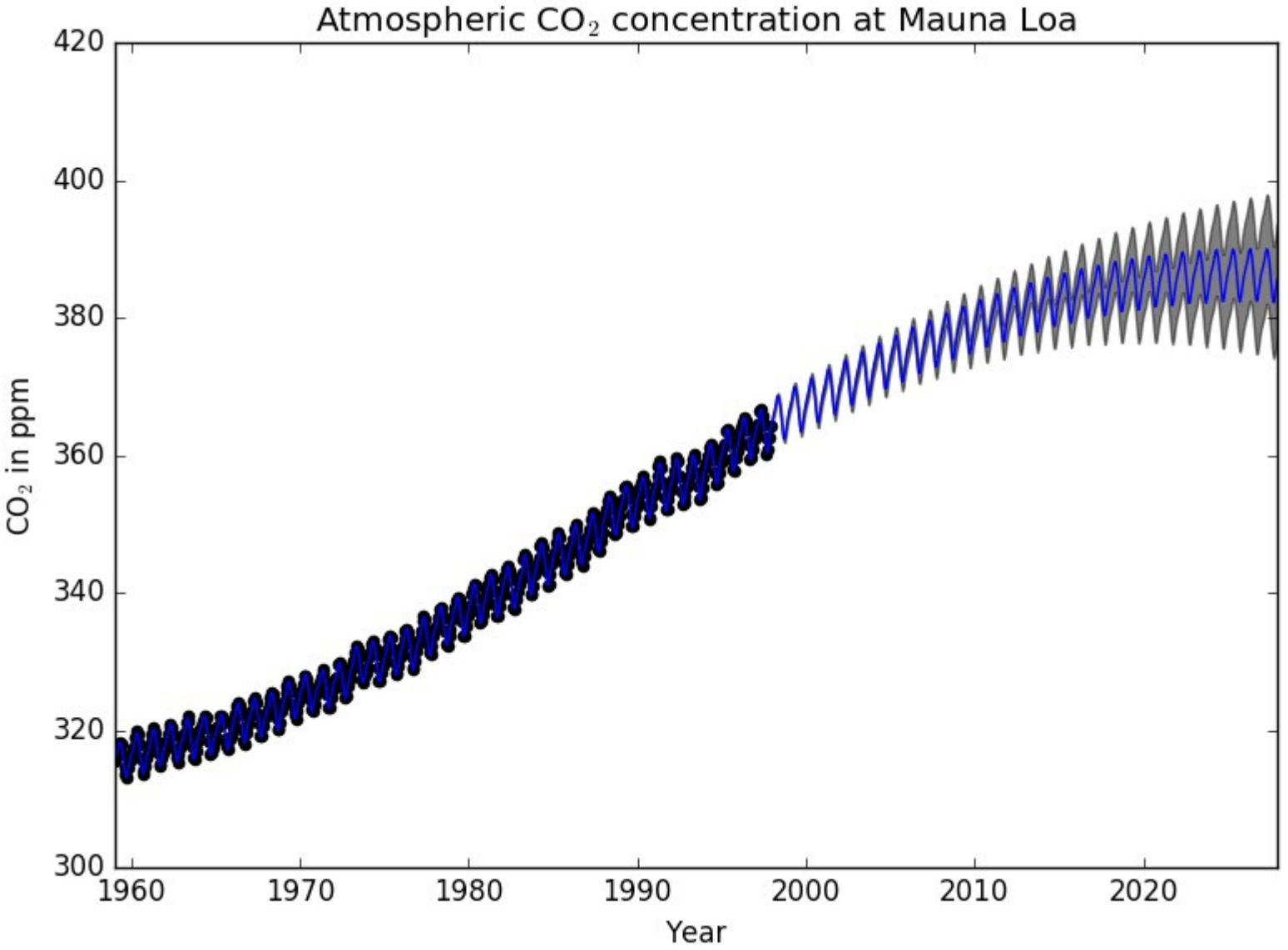
```
from sklearn.model_selection import KFold
cv = KFold(n_folds)
for train, test in cv.split(X, y):
    ...
```

```
from sklearn.mixture import GaussianMixture
from sklearn.mixture import BayesianGaussianMixture
```

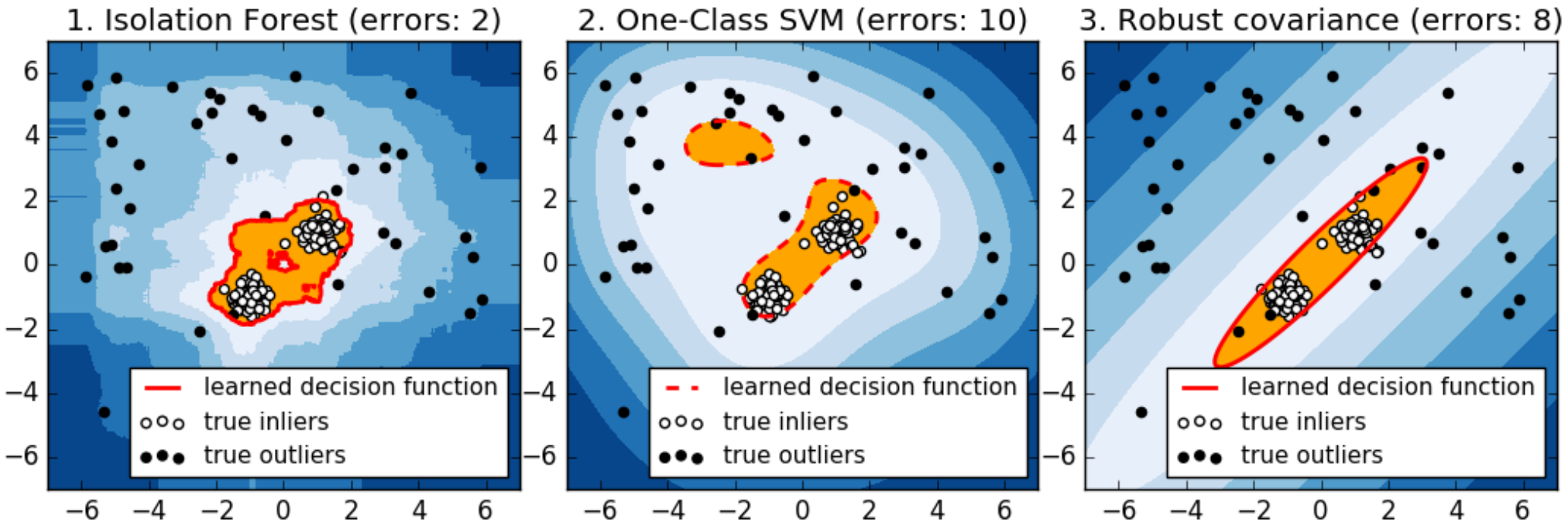



PCA()
RandomizedPCA()
PCA()

Gaussian Process Rewrite



Isolation Forests



Play

```
from sklearn.neural_network import MLPClassifier
```

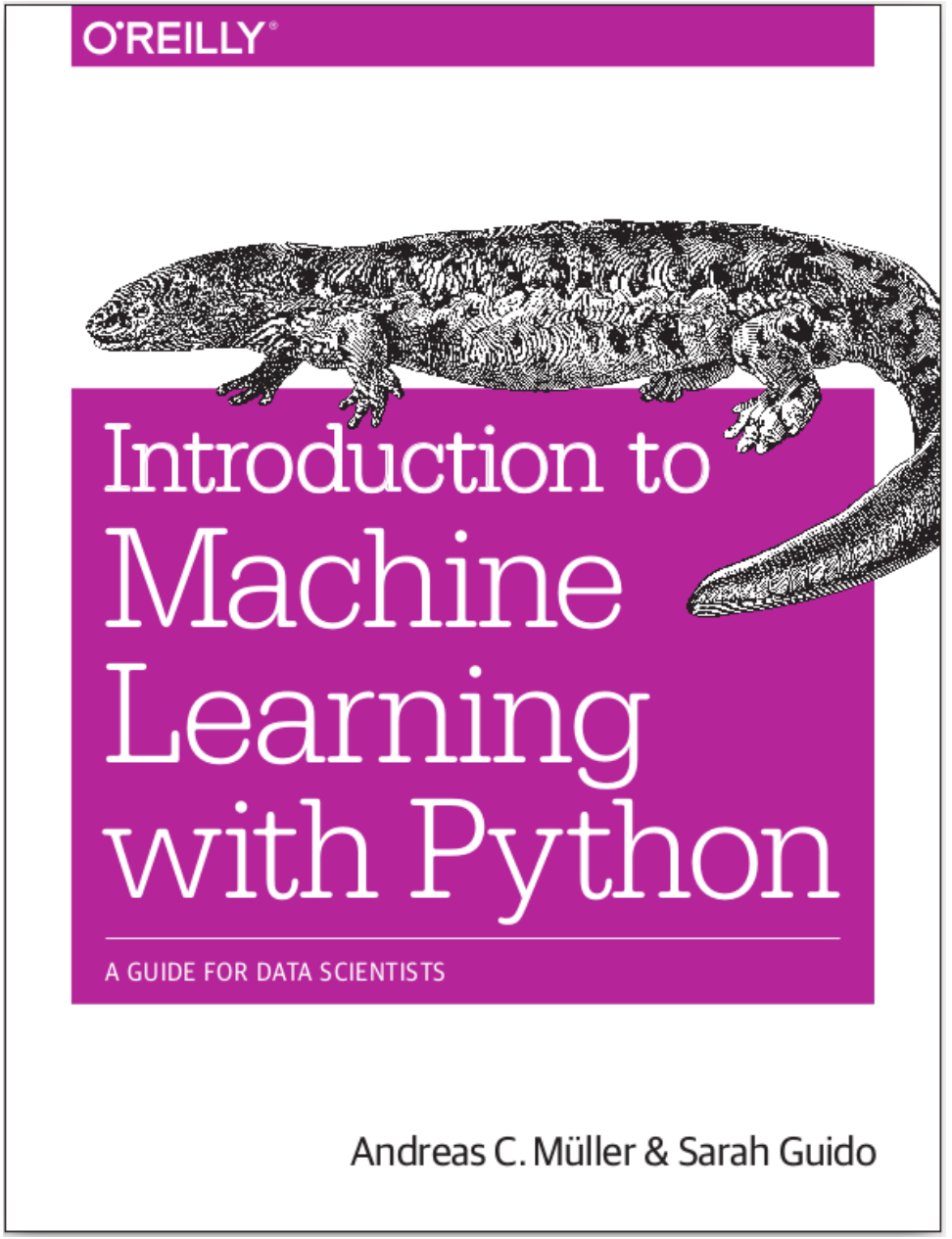
Work

```
import keras
```

```
pipe = Pipeline([('preprocessing', StandardScaler()),
                 ('classifier', SVC())])

param_grid = {'preprocessing': [StandardScaler(), None]}

grid = GridSearchCV(pipe, param_grid)
```


O'REILLY
Introduction to
Machine
Learning
with Python
A GUIDE FOR DATA SCIENTISTS
Andreas C. Müller & Sarah Guido

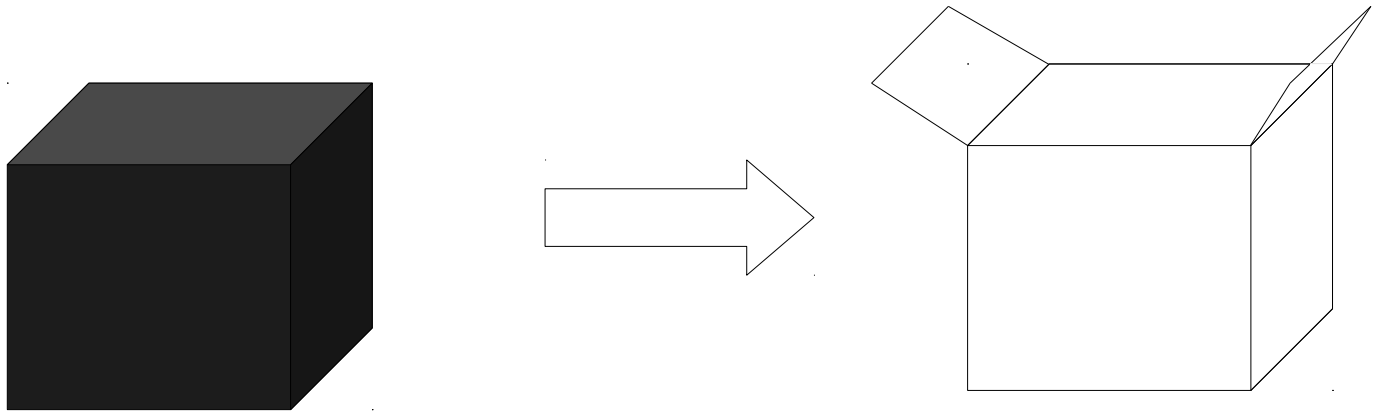
(further) Future

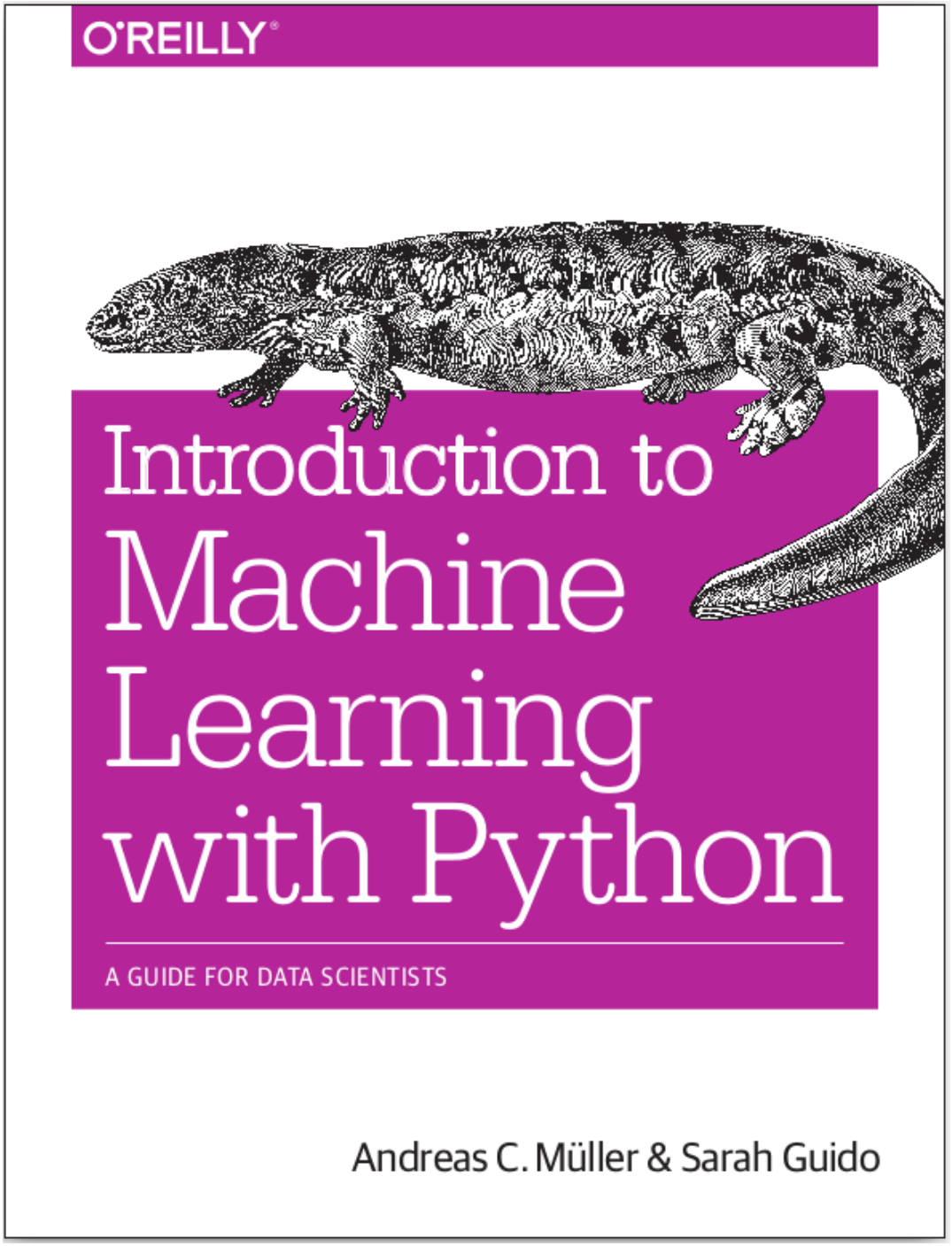
Feature / Column names

```
from __future__ import sklearn.plotting
```

```
from __future__ import AutoClassifier
```

More Transparency





amueller.github.io

@amuellerml

@amueller

t3kcit@gmail.com